

RESEARCH OUTPUTS / RÉSULTATS DE RECHERCHE

La robotisation de la vie ou la tentation de l'inséparation

Rouvroy, Antoinette

Published in:

Intelligence artificielle et droit

Publication date:

2017

Document Version

le PDF de l'éditeur

[Link to publication](#)

Citation for pulished version (HARVARD):

Rouvroy, A 2017, La robotisation de la vie ou la tentation de l'inséparation. Dans *Intelligence artificielle et droit*. Collection du CRIDS, Numéro 41, Larcier , Bruxelles, p. 13-46.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

TITRE 1

La robotisation de la vie ou la tentation de l'inséparation

Antoinette ROUVROY*

« There was a time when humanity faced the universe alone and without a friend. Now he has creatures to help him ; stronger creatures than himself, more faithful, more useful, and absolutely devoted to him. Mankind is no longer alone ».

(Isaac Asimov, *I, Robot*, 1950)

Introduction

Le cours médiatique de l'intelligence artificielle ne s'est jamais mieux porté. Les robots « intelligents », voir supra-« intelligents », seraient en approche. La victoire, en début d'année 2017, de l'algorithme Libratus développé à l'Université de Carnegie Mellon contre certains des meilleurs joueurs de poker mondiaux marque un jalon dans le développement de l'intelligence artificielle, démontrant la capacité d'un algorithme à surpasser les êtres humains en situation d'information incomplète. « On peut imaginer que cette technologie pourrait déjouer les marchés financiers, surpasser les chercheurs humains en matière d'inventions, manipuler les dirigeants humains, et développer des armes que nous ne pouvons même pas comprendre », avertissait déjà Stephen Hawking en 2014¹. Assistants personnels, *chatbots* ou agents conversationnels quasi indiscernables des

* Chercheur qualifié du FRS-FNRS au Centre de Recherche Information, droit et Société, Université de Namur. Certaines parties du présent chapitre sont fondées sur le rapport « des données et des hommes : droits et libertés fondamentaux dans un monde de données massives » rédigé par l'auteure à destination du Conseil de l'Europe, T-PD-BUR(2015)09REV.

¹ St. HAWKING, « Transcendence looks at the implications of artificial intelligence – but are we taking AI seriously enough? », *The Independent*, 1^{er} mai 2014. Voy. égal. N. BOSTROM, *Superintelligence: Paths, Dangers, Strategies*, Audible Studios on Brilliance Audio, MP3 Una edition, 2015.

usagers humains sur les réseaux sociaux, systèmes automatisés de recommandation et d'aides à la décision, voitures autonomes, robots pédagogiques dans les écoles, robots gériatriques dans les maisons de retraite, robots compositeurs², robots scénaristes³, robots-chirurgiens, avocats et juges robots, drones armés autonomes. Il ne se trouverait plus aucune activité humaine – pas même celles qui consistent à juger son prochain ou à lui ôter la vie –, qui ne semble promise au grand remplacement des Humains par les machines à en croire les récits robot-utopiques ou robot-apocalyptiques qui parcourent l'univers des médias mis en concurrence pour capter l'attention raréfiée d'une audience versatile.

Le remplacement total du travail humain par les robots pourrait advenir, nous dit-on, endéans les 120 prochaines années⁴. D'après des « experts en IA », les machines auront surpassé les humains dans les tâches de traduction linguistique en 2024, dans la rédaction de dissertations en 2031, dans la conduite de camions en 2027, dans la vente en 2031. D'ici 45 ans, il y aurait 50 % de chances pour que les intelligences artificielles surpassent les êtres humains dans la totalité des « tâches », ce qui devrait aboutir, d'ici 120 ans, au grand remplacement⁵. Le 19 janvier 2017, Laurent Alexandre, dans une intervention qu'il faisait devant le Sénat français, augurait que « dans le futur, tous les gens qui ne seront pas complémentaires de l'intelligence artificielle seront soit au chômage, soit avec un emploi aidé. Pas 99 % des gens. 100 %. (...) Aucun emploi non complémentaire de l'IA n'existera en 2050. Par définition ».

Force est de constater que, pour l'heure, le tournant numérique a surtout donné un grand coup d'accélérateur aux processus de flexibilisation (ou de précarisation) du travail et de contournement des défenses corporatistes ou syndicales, et à l'accroissement considérable des capacités de surveillance continue de la productivité des travailleurs mis en concurrence à l'échelle quasi-moléculaire de la donnée, le tout sous les traits avenants de la *gamification*, de l'*uberisation*, de la *contributive economy* ou de la *sharing economy*,... On parle de *disruption*, de révolution ou de

² http://www.lemonde.fr/les-cles-de-demain/article/2016/12/01/not-easy-un-single-signé-watson_5041630_4758288.html

³ http://www.lemonde.fr/pixels/article/2016/06/10/une-intelligence-artificielle-écrit-le-scénario-d-un-court-métrage_4947819_4408996.html

⁴ K. GRACE et al., « When Will AI Exceed Human Performances? Evidence from AI Experts », *ArXiv*, 30 mai 2017 (<https://arxiv.org/pdf/1705.08807.pdf>). Voy. aussi B. STIEGLER, *La société automatique 1. L'avenir du travail*, Paris, Fayard, 2015. Voy. E. OSNOS, « Doomsday prep for the super-rich », *The New Yorker*, 30 janvier 2017. <https://www.newyorker.com/magazine/2017/01/30/doomsday-prep-for-the-super-rich>.

⁵ Voy. aussi C. B. FREY and M. A. OSBORNE, « Technology at work: The future of innovation and employment », *Citi GPS: Global Perspectives & Solutions*, February 2015. http://www.oxfordmartin.ox.ac.uk/downloads/reports/Citi_GPS_Technology_Work.pdf

transition numérique comme s'il s'agissait à présent de quitter l'ancien monde hérité des révolutions industrielles, avec ses protections sociales, ses représentations syndicales, ses assurances mutuelles, pour entrer dans un monde radicalement neuf, le monde hyper flexible, hyper liquide, hyper individualiste et darwiniste des start-ups.

À côté de cette flexibilisation du travail et de la vie (les deux devenant indistincts dans le capitalisme 24/7 que décrit Jonathan Crary⁶) d'une classe moyenne hyper-prolétarisée se développe l'utopie aristocratique des transhumanistes : la perspective d'une « augmentation » ou d'un perfectionnement des performances et de la santé des êtres humains à travers des formes plus ou moins intégrées de symbiose avec les machines à la faveur d'une convergence entre nanotechnologies, biotechnologies, technologies de l'information, et sciences cognitives (NBIC). Cette « singularité » n'étant destinée qu'à ceux qui pourront en payer le prix, les promoteurs d'une humanité augmentée font majoritairement partie d'une oligarchie extrêmement minoritaire de super-riches dont certains sont disposés à se désolidariser du restant de l'espèce humaine, ceux-là-même qui, en bons *survivalistes*, investissent dans des îles désertes et des bunkers de luxe afin d'être prêts pour l'apocalypse – celle-ci pouvant prendre la forme d'une guerre nucléaire, d'une catastrophe écologique ou de révoltes sociales⁷.

Quoi qu'il en soit, la fascination contemporaine pour l'intelligence artificielle et pour les perspectives utopiques ou dystopiques de la « singularité » technologique popularisées par les techno-prophètes de la *Silicon Valley* et leurs alliés académiques produit – délibérément ou non – des effets de sidération qui empêchent de percevoir les défis que nous lance le présent⁸. Il y a, autour de l'intelligence artificielle, tout un art du *buzz*, qui distrait d'une série d'enjeux moins spectaculaires que la perspective de robots super-intelligents menaçant l'humanité d'obsolescence voir d'extinction, mais néanmoins absolument actuels et cruciaux : notre prédisposition à sous-traiter la charge de décider à des machines; la concentration du pouvoir entre les mains des concepteurs et des détenteurs des plus grandes masses de données ; les innombrables « biais » que transposent et « naturalisent » les données sous la forme univoque du langage binaire ; les « biais » des algorithmes eux-mêmes ; la propension des algorithmes, aussi « objectifs » soient-ils, à prendre des « bruits » pour des « signaux » dès lors que la masse des données s'accroît ; la tendance contemporaine à nous satisfaire de corrélations sans plus chercher à établir les causes ; ...

⁶ J. CRARY, *Late Capitalism and the Ends of Sleep*, Verso, 2014.

⁷ E. OSNOS, « Doomsday Prep for the Super-Rich », *The New Yorker*, 30 janvier 2017.

⁸ Pour une critique de ces mythes technologiques, voy. J.-G. GANASCIA, *Le mythe de la Singularité. Faut-il craindre l'intelligence artificielle ?*, coll. Science ouverte, Paris, Seuil, 2017.

Le nom « robot », ou la notion d'Intelligence Artificielle recouvrent un ensemble trop vaste et diversifié de phénomènes et d'objets techniques pour tous les réduire à un commun dénominateur, à part celui-ci : tous absolument tributaires du « tournant computationnel » dans lequel nous sommes tous bel et bien engagés *dans le présent* : une « mise en données » du monde et de ses habitants d'une ampleur sans précédents⁹ – les *Big Data* – et une montée en puissance du « paradigme indiciaire » – faisant des données, mêmes les plus triviales, les moins personnelles, les nouvelles coordonnées de modélisation du monde, en ce compris le monde social¹⁰. Pour les interfaces cognitives fondées sur les technologies numériques, l'unité de perception, de compréhension du monde n'est pas la phrase, le mot, le signe, toujours porteurs de significations, mais la donnée numérique, signal a-signifiant mais calculable.

Ce tournant computationnel n'est pas tant une révolution qu'une « réplique » plus intense de séismes épistémiques antérieurs. La prolifération de données en quantités massives n'est pas sans rappeler « l'avalanche de nombres imprimés »¹¹ et la révolution probabiliste provoquée par l'explosion des statistiques numériques en Europe entre (*grosso modo*) 1820 et 1840.¹² Quant au paradigme indiciaire, « apparu silencieusement, à la fin du XIX^e siècle, dans le domaine des sciences humaines (...) un modèle épistémologique (ou, si l'on préfère, un paradigme) » – en particulier, dans le domaine de l'enquête criminelle, de la psychanalyse, et de l'histoire de l'art – il consiste en ceci que « [D]es traces parfois infinitésimales permettent d'appréhender une réalité plus profonde. Des traces, plus précisément des symptômes (dans le cas de Freud), des indices (dans le cas de Sherlock Holmes), des signes picturaux (dans celui de Morelli) »¹³.

⁹ Il y avait bien, avant l'Internet, de la recherche en intelligence artificielle, mais la disponibilité de données en grandes quantités permise par l'Internet a provoqué une accélération importante dans le développement de ces recherches et leur déploiement en dehors de systèmes fermés.

¹⁰ Par la grâce des algorithmes, c'est à partir des données numériques, plutôt qu'à partir de processus institutionnels ou délibératifs, que se construisent les catégories à travers lesquelles les individus sont classés, évalués, récompensés ou sanctionnés, ou à travers lesquelles s'évaluent les mérites et les besoins des personnes ou encore les opportunités ou la dangerosité que recèlent les diverses formes de vie qu'elles habitent. Voy. à ce propos A. ROUVROY, « Gouverner hors les normes : la gouvernementalité algorithmique », *Lacan Quotidien*, n° 733, 6 juillet 2017.

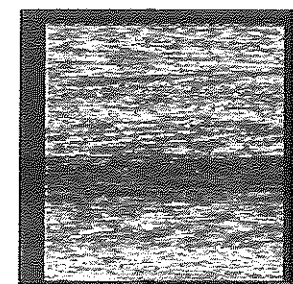
¹¹ I. HACKING, *The taming of chance*, Cambridge, Cambridge University Press, 1990, p. 18.

¹² M. LETA (Ambrose) JONES, « Lessons from the Avalanche of Numbers: Big Data in Historical Context » (August 1, 2014), *I/S: A Journal of Law and Policy for the Information Society*, 2014-2015. SSRN : <https://ssrn.com/abstract=2486981>.

¹³ C. GINZBURG, « Signes, traces, pistes. Racines d'un paradigme de l'indice », *Le débat*, n° 6, 2010.

Dans le paradigme indiciaire contemporain dominé par les techniques de *patterns recognition*, ou la détection automatique de corrélations au sein des masses de données, « voir » est supplanté par « détecter » ou plutôt « faire détecter » et « pré-voir » par les algorithmes. En témoigne, mieux que tout discours, une image classifiée « secrète » extraite des dossiers détenus par la NSA révélés par Edward Snowden : on n'y voit qu'un brouillard de pixels de différentes nuances de gris. Aucun risque, donc, que la seule vue de cette image puisse révéler quoi que ce soit à qui que ce soit. La vision de ce qui est dans le présent perd en importance face au filtrage, au décryptage et à la détection de corrélations (*pattern recognition*) « prédictives » de ce qui pourrait advenir¹⁴. La « vision » algorithmique, comme une sorte d'échographie du futur, rend appréhendable ce qui n'est pas (encore) présent à la vue et produit, de ce fait, de nouveaux « objets », de nouveaux « espaces » – des objets absolument infigurables – pour l'optimisation et la préemption¹⁵. Il faut bien comprendre que ces nouveaux objets spéculatifs ne sont pas des choses – il n'y a pas de choses, d'objets résilients dans l'univers algorithmique : seuls existent des réseaux de données évolutifs en temps réel, des profils, des « patterns », qui ne sont rien d'autre que des agrégats temporaires de données. Il faut bien comprendre aussi que ces « motifs », « profils », « modèles » ou « patterns » produits par les algorithmes dits prédictifs ne sont ni vrais, ni faux. Ils produisent tout au plus des effets de fiabilité sans vérité : une fiabilité suffisante pour nourrir des stratégies d'action. Ils permettent d'agir par avance sur ce qui n'existe, au moment de la détection, que sur le mode de la possibilité, ou de la probabilité.

SECRET



¹⁴ H. STEYERL, « A Sea of Data : Apophenia and Pattern (Mis-)Recognition », *e-flux*, #72, avril 2016.

¹⁵ À ce propos, voy. O. ALPERN, *Beautiful Data. A History of Vision and Reason since 1945*, Duke University Press, 2014.

D'une civilisation du texte et du signe nous serions en passe de basculer vers une civilisation de l'algorithme et du signal. Voilà qui, en soi, est proprement bouleversant. Il s'agirait, aux dires de certains, d'une nouvelle étape dans l'histoire des formes d'écriture. Après tout, Alan Turing lui-même reconnaissait que « mécanisme » et « écriture » sont *grosso modo* synonymes¹⁶. C'est en tant que figures concrètes de ce basculement sémiotique et épistémologique plutôt qu'en tant que simulacres phantasmatiques de l'intelligence humaine, que je m'intéresserai ici aux robots et à l'IA.

Ce basculement sémiotique et épistémologique affecte bien évidemment très fondamentalement les modes de production du « savoir », certainement dans les sciences biologiques et les sciences de la nature (astronomie, climatologie, génétique, ...) mais également les sciences humaines (*digital humanities*, ...). Il affecte aussi, de façon tout à fait cruciale, avec des conséquences rien moins que triviales, les modes d'exercice du pouvoir ou de « gouvernement » des conduites.

Ainsi, dans l'hypothèse d'une gouvernementalité algorithmique¹⁷ c'est à partir des signaux numériques transpirant des trajectoires, relations et interactions des individus, plutôt qu'en fonction de normes émanant de processus délibératifs antérieurs que se construisent les profils/scores/appariements à travers lesquels les individus sont classés, évalués, récompensés ou sanctionnés, à travers lesquels s'évaluent les mérites et les besoins ou encore les opportunités ou la dangerosité que recèlent les formes de vie. La « vision » algorithmique « défait » nos « images de pensée » (les catégories, les préjugés à travers lesquelles nous sommes pré-disposés à percevoir et à évaluer les événements du monde et ses habitants), au double sens où, d'une part, elle les fragmente sous forme de données numériques asignifiantes mais métabolisables par les systèmes informatique, et, d'autre part, elle disqualifie les formes que nous, êtres humains, pouvons « reconnaître », nos images qui, face à l'« objectivité machinique », paraissent trop peu fiables pour légitimer aucun savoir, aucune prédiction.

¹⁶ Voy., par exemple, les travaux de J. LASSÈGUE, <http://www.formes-symboliques.org/jean-lassegue/>

¹⁷ Sur ce thème de la gouvernementalité algorithmique, nous nous permettons de renvoyer le lecteur à nos travaux précédents, notamment A. ROUVROY et Th. BERNIS, « Gouvernementalité algorithmique et perspectives d'émancipation. Le disparate comme condition d'individuation par la relation ? », *Réseaux*, vol. 177, n° 1, 2013, pp. 163-196 ; A. ROUVROY et Th. BERNIS, « Le nouveau pouvoir statistique. Ou quand le contrôle s'exerce sur un réel normé, docile et sans événement car constitué de corps "numériques" », *Multitudes*, vol. 40, n° 1, 2010, pp. 88-103.

CHAPITRE 1. De quoi la robotisation de la vie est-elle le nom ?

Reprenons la citation d'I. Asimov, extraite de son roman visionnaire, *I, Robot* : « Il fut un temps où les hommes se tenaient seuls et sans amis devant l'univers. À présent ils ont des créatures pour les assister ; des créatures plus fortes, plus fiables, plus utiles et qui lui sont absolument dévouées. L'humanité n'est plus seule ».

Le « tournant computationnel », ou le déploiement intensif d'une multitude d'appareils électroniques fondés sur la technologie numérique, véritables prothèses cognitives, mémorielles, affectives, communicationnelles capables, à notre place, de détecter les événements, transformations, altérations subtiles du milieu et d'y « répondre » pour nous, à notre place, nous immerge dans l'environnement contrôlé-personnalisé de la réalité numérique et nous immuniseraient de l'incertitude radicale à laquelle nous sommes exposés, « seuls face à l'univers ».

Sans doute la source énigmatique de notre engouement pour les robots et pour l'intelligence artificielle provient-elle effectivement de cette angoisse de *séparation* que les robots viennent soulager. C'est une véritable passion que vouent les individus à leurs objets connectés, certains avouant qu'ils se sentent « nus » ou ont l'impression d'être privés d'une partie de leur vie sans leur smartphones et autres « doudous » pour adultes et adolescents. Il arrive même qu'en Chine des adolescents vendent un rein afin de pouvoir s'acheter le dernier modèle d'iPhone¹⁸. « L'affection que les gens portent aux ordinateurs se transfère sur tous les objets, même les plus ordinaires. Et plus les gens aiment utiliser l'ordinateur pour tout, plus la vie semble être incomplète si elle n'a pas de liens avec l'informatique »¹⁹. Nous dispensant du même coup des artifices symboliques, politiques, institutionnels de la représentation, nous dispensant d'avoir à assumer l'invention individuelle et collective du « sens » à travers laquelle, tant bien que mal, nous conjurons, – les légitimant par la possibilité toujours ouverte de la contradiction herméneutique²⁰ et de la

¹⁸ J. QIU, « Goodbye ISlave: Rethinking Smartphone, Activism, and Chinese Labor », *6th ICTs and Society Conference: Digital Objects, Digital Subjects, An Interdisciplinary Symposium on Activism, Research & Critique in the Age of Big Data Capitalism*, Londres, University of Westminster, 20-21 mai 2017, <https://www.youtube.com/watch?v=eTPlBiVZb70>

¹⁹ I. BOGOST, « You Are Already Living Inside a Computer. Futurists predict a rapture of machines, but reality beat them to it by turning computing into a way of life. », *The Atlantic*, 14 septembre 2017.

²⁰ Ce dont il s'agit, c'est de « voir à l'intérieur de la vue », selon la formule de Max Ernst, ou de regarder comme on voit.

critique²¹ – les effets angoissants de notre « solitude face à l'univers », la robotisation de la vie, c'est l'utopie d'une immanence totale, remplaçant les points-de-vue séparés et partiels, les représentations inadéquates, le retard de nos représentations par une juxtaposition de « maintenant » successifs, dans lesquels les choses parleraient d'elles-mêmes, une « rechosification » généralisée dans un langage univoque de 1 et de 0.

Pierre Legendre énonçait l'impossibilité dans laquelle se trouve l'Homme de rejoindre l'univers autrement qu'à travers ses représentations :

« [L]e monde n'est pas donné à l'homme, si ce n'est par le langage qui le sépare des choses et le divise lui-même. Et cette déchoséification généralisée, une sorte de dématérialisation de la matérialité du monde, lui impose le joug d'un univers de la représentation, su et insu, bien plus difficile d'accès à nos investigations que celui où se meuvent, pour survivre elles aussi, les autres espèces perfectionnées »²².

L'idéologie technique des *big data* ne consiste-t-elle pas, précisément, à prétendre que le « sens », comme le « savoir », n'a plus à être produit puisqu'il se trouverait toujours déjà là : dans les masses de données sur lesquelles il « suffirait » de faire tourner des algorithmes dont l'intelligence consiste, précisément, à être capables de détecter, à la vitesse de la lumière, des corrélations suffisamment « significatives » pour dispenser de toute interprétation, de toute théorie ? Dans l'univers numérique qui, progressivement, nous tient lieu de monde, opérant dans un décor numérique artificiellement éternel et illimité, informe donc, qu'elle tend à faire passer pour notre monde, la gouvernementalité algorithmique participe d'un *régime d'inséparation* :

« Un mouvement de fond nous a fait passer d'un univers humaniste composé d'entités séparables à un réel inséparable où tous les phénomènes devenus globalisés sont liés, en interrelation et en co-dépendance. Des smartphones au multiculturalisme, des défis de l'écologie à la politique de réinvention des frontières, du politically correct au posthumanisme, notre monde élabore une nouvelle condition d'existence, dont la figure de l'Autre a disparu. Est-ce bien ? Est-ce mal ? Là n'est pas la question.

²¹ Par « critique », nous entendons « une pratique qui suspend le jugement et une occasion pour de nouvelles pratiques de valeurs fondée précisément sur cette suspension. Dans cette perspective, la critique s'attache à la constitution occlusive du champ des catégories elles-mêmes plutôt qu'à la subsumption du cas particulier sous une catégorie préconstituée » (M. FOUCAULT, « Qu'est-ce que la critique ? », Compte rendu de la séance du 27 mai 1978, *Bulletin de la Société française de Philosophie*, avril-juin 1990).

²² P. LEGENDRE, *De la société comme texte. Linéaments d'une anthropologie dogmatique*, Paris, Fayard, 2001, p. 17.

Cela est. Ce monde enfanté par l'inséparation, la politique, les sciences, les arts et les techniques en ont produit l'architecture – une architecture sans coupure ni soudure »²³.

Deep Mind filiale de *Google* chargée des projets d'intelligence artificielle annonçait tout récemment s'être dotée d'une nouvelle « AI ethics unit », intitulée *DeepMind Ethics & Society* (DMES), consacrée exclusivement aux enjeux sociétaux des perspectives ouvertes par les progrès de l'intelligence artificielle. La « vision » conférant son impulsion à la DMES et justifiant le traitement « en interne » de ces enjeux sociétaux (*in-housing*) – plutôt, par exemple, que la participation de représentants de l'entreprise à des débats publics relatifs à l'éthique de l'IA – est celle d'un dépassement des approches sectorielles. Selon le co-fondateur de *Deep Mind*, Mustafa Suleyman : « Si nous voulons être gagnants comme civilisation dans les prochaines dix, vingt ou trente années, il est absolument crucial de dépasser le mode de pensée par défaut, qui est un mode de pensée sectoriel, dans lequel les académiques reprochent au gouvernement de ne pas les écouter, où le gouvernement reproche aux entreprises de ne pas comprendre ses perspectives, où les entreprises rejettent les arguments des deux premiers. On est enfermés dans ce cercle vicieux où chacun joue son rôle institutionnel et rejette les autres. Cela doit cesser. »²⁴ La problématique qui surgit inévitablement à l'égard de cette forme d'accaparement ou de mise sous tutelle, par une entreprise privée de la puissance de *Google*, du débat éthique autour de l'intelligence artificielle, indépendamment des (bonnes) intentions de *Google*, est celle du court-circuitage anticipatif et de facto de tout débat législatif. Or il n'est pas évident que les enjeux disruptifs de l'intelligence artificielle et l'« oeucuménisme » institutionnel-industriel-scientifique ici prôné justifie la mise hors jeu du débat démocratique et de sa formalisation juridique, en ce compris les « coupures » et « soudures » que le droit, dans toute son artificialité, impose en vertu de conceptions de la justice, notamment, éminemment tributaires du type de théorie de la justice ou de contrat social qui « fait tenir ensemble » la société des individus.

²³ D. QUESSADA, *L'inséparable, essai sur le monde sans Autre*, coll. « Perspectives critiques », Paris, PUF, 2013.

²⁴ J. TEMPERTON, « DeepMind's new AI ethics unit is the company's next big move », *Wired*, 4 octobre 2017.

CHAPITRE 2. Vices et vertus de l'objectivité machinique

Dans un premier temps, les attentes générées par le phénomène des *Big Data* convergent vers la perspective d'amélioration (dans le sens d'une objectivation et d'une optimisation) des décisions dans une multitude de secteurs : sécurité et prévention du terrorisme, optimisation et distribution de la présence policière, soins de santé, politiques énergétiques, gestion du trafic et optimisation des transports en commun, prévention des fraudes, marketing et amélioration de l'« expérience des utilisateurs ou des consommateurs », différenciation des prix en fonction des profils des consommateurs, gestion des stocks, gestion du trafic, orientation en éducation et formation, recrutement et gestion des ressources humaines...

Les *Big Data* induisent par exemple de nouvelles attentes en termes de planification « objective » et « réactive en temps réel » des politiques publiques. Collectées notamment à travers tous les capteurs impliqués dans le déploiement du concept de ville intelligente et à travers les appareils de téléphonie mobile, les *Big Data* promettent de pouvoir prendre la mesure objective, en temps réel, de la vie urbaine et de ses infrastructures, au profit d'un développement, d'une gestion, d'une réglementation et d'une vie dans la ville fondées sur les données – donc sur une forme numérique, quantitative d'évidence rationnelle. Les *Big Data* pourraient par exemple permettre d'optimiser la fréquence, les horaires et les trajectoires des véhicules de transports en commun en fonction des intérêts collectifs déduits de la géolocalisation.

Cependant, les nouvelles capacités de surveillance numérique, d'analyse préemptive et de décision automatisée accompagnant le tournant computationnel ravivent aussi, crucialement, la question fondamentale de la définition du pacte social existant entre les individus, les entreprises et les États. La présentation des enjeux en termes d'innovation, de compétitivité, d'intérêts individuels des consommateurs ou des utilisateurs occulte bien souvent les enjeux éthiques, juridiques, politiques du tournant computationnel, au risque de porter atteinte à la prééminence du droit ainsi que des droits de l'homme et des libertés fondamentales²⁵. Pourtant, les

²⁵ Nous nous permettons de renvoyer le lecteur à A. ROUVROY, *Des données et des Hommes. Droits et Libertés fondamentaux dans un monde de données massives*, Rapport au Conseil de l'Europe, janvier 2016 (T-PD-BUR(2015)09REV Ed.). Version augmentée disponible : https://www.academia.edu/15730565/DES_DONN%C3%89ES_ET_DES_HOMMES._DROITS_ET_LIBERT%C3%89S_FONDAMENTAUX_DANS_UN_MONDE_DE_DONN%C3%89ES_MASSIVES_version_augment%C3%A9e_

capacités nouvelles fondées sur « l'intelligence des données », dont beaucoup restent imperceptibles ou inaccessibles au citoyen, peuvent aggraver considérablement l'asymétrie d'information et/ou de pouvoirs entre ceux qui détiennent les données, ceux qui conçoivent les algorithmes et ceux qui voient leurs possibilités et opportunités façonnées en fonction de ces données transpirant de leurs comportements, actions et interactions.

Pour autant, si, sur un mode quasi-réflex, les citoyens et les juristes voient dans la « transparence numérique » postulée par l'objectivité machinique et l'opérationnalité des dispositifs algorithmiques nourris par les *Big Data* une menace pour la vie privée des citoyens à conjurer prioritairement à travers les régimes de protection des données personnelles, il n'est absolument pas évident du tout que le prisme de la protection des données personnelles soit le plus adéquat pour faire face à ce qui menace concrètement les droits et libertés fondamentaux à l'ère des données massives. Le phénomène des données massives met d'ailleurs les régimes juridiques de protection des données personnelles « en crise ». La prolifération et le traitement de quantités massives de données – présumées par le phénomène des *Big Data* et par les solutions robotiques qui s'en nourrissent – entrent en opposition frontale avec les grands principes de la protection des données : la *minimisation* (on ne collecte que les données nécessaires au but poursuivi), la *finalité* (on ne collecte les données qu'en vue d'un but identifié, déclaré, légitime), la *limitation dans le temps* (les données doivent être effacées une fois le but atteint, et ne peuvent être utilisées, sauf exceptions, à d'autres fins que les fins initialement déclarées). Les *Big data*, c'est au contraire une collecte maximale, automatique, par défaut, et la conservation illimitée de tout ce qui existe sous une forme numérique, sans qu'il y ait, nécessairement, de finalité établie a priori puisque l'utilité des données ne se manifeste qu'en cours de route, à la faveur des pratiques statistiques de datamining, de machine-learning, etc.

Par ailleurs, la focalisation des régimes de protection sur les seules données personnelles et l'individualisme méthodologique sous-tendant l'exigence du consentement individuel – une exigence intéressante en cela qu'elle rend les personnes individuelles sensibles à la nécessité de faire des choix, mais qui ne permettant pas de faire face aux enjeux collectifs – les rend insensibles au fait qu'aujourd'hui, ce qui rend problématique ou sensible le traitement de telle ou telle donnée n'est pas « dans » cette donnée mais dépend de la force de la corrélation existant « entre » cette donnée et d'autres données recueillies par ailleurs. C'est le paradoxe de la « privacy-interdependency » – un paradoxe qui était déjà évident dans le contexte de la recherche en génétique humaine mais qui, données massives aidant, contamine littéralement tous les domaines – consistant en ceci que les

choix de divulgation de « ses » données personnelles par mon voisin permettent d'inférer, à mon propos, des informations de type probabiliste mais sur base desquelles des décisions peuvent être prises à mon égard. Ce qui m'est véritablement personnel et singulier perd en importance prédictive face aux inférences produites par l'analyse des données relatives aux personnes avec qui je suis en relation, ou avec qui je partage un certain nombre d'éléments a priori triviaux comme le fait de faire mes courses dans le même supermarché qu'elles, ou d'habiter le même quartier. C'est une image indifférente aux biographies personnelles, une image « épidémique » du risque de non remboursement de crédit, du risque de fraude, du risque d'absentéisme,... que produisent les algorithmes à l'heure où nous glissons d'une société de l'assurance à une société du « scoring ».

Si l'on se place du point-de-vue de l'individu, d'ailleurs, le problème n'est pas celui d'une plus grande visibilité sociale ni d'une disparition de la sphère privée. On assiste au contraire à une hypertrophie de la sphère privée – une privatisation amniotique des espaces publics dans lesquels chacun évolue, contenu dans sa « bulle » connectée/personnalisée – au détriment de l'espace public comme espace où l'on est exposé à ce qui n'a pas été prévu pour nous. On ne peut pas même affirmer à coup sûr que les technologies contemporaines de l'information et de la communication nous rendent vraiment plus « visibles ». Les « demoiselles du téléphone » de jadis, entremetteuses incontournables et pas toujours discrètes des rendez-vous galants dans des microcosmes sociaux avides de rumeurs, représentaient une menace au moins aussi importante pour la protection des données personnelles et de la vie privée des personnes que les algorithmes aveugles et sourds des moteurs de recherche d'aujourd'hui. Peut-être n'avons-nous jamais été moins « visibles », moins « signifiants » dans l'espace public en tant que sujets, en tant que personnes, qu'aujourd'hui. La prolifération des selfies et autres performances identitaires numériques est symptomatique à cet égard. L'incertitude d'exister induit une pulsion d'édition de soi sans précédent : se faire voir pour croire en sa propre existence.

Nous vivons, objectera-t-on, sous le regard continu, ubiquitaire, des algorithmes. Mais en quoi ce regard sans yeux des algorithmes menace-t-il prioritairement la vie privée des citoyens? Est-ce vraiment *d'abord* un problème de protection de la vie privée? Ce qui intéresse les bureaucraties privées et publiques qui nous « gouvernent », – et en vue de quoi les algorithmes sont paramétrés – c'est de détecter automatiquement, sans nécessairement nous identifier ni nous interpeller, nos « potentialités », nos propensions, ce que nous pourrions désirer, ce que nous serions capables de commettre, avant que nous en soyons nous-mêmes conscients. Une propension, un risque, une potentialité, ce n'est pas encore une personne,

et les inférences fondées sur des corrélations entre données numériques ne sont pas des faits. Les phéromones numériques parlent de et pour nous mais ne disent pas qui nous sommes : ils disent ce dont nous sommes capables. L'enjeu, ce n'est pas tant la donnée personnelle que la disparition de la « personne » dans les deux sens du terme. Il nous devient impossible de n'être « personne », d'être « absents » (nous ne pouvons pas ne pas laisser de traces) et il nous est impossible de compter en tant que « personne », anticipés comme nous sommes par les profilages, « scorings », appariements algorithmiques. Ce que nous pourrions dire de nous mêmes ne devient-il pas redondant, sinon suspect, face à l'efficacité et à l'objectivité machinique des traitements automatiques dont nos « data-doubles » font l'objet?

Cependant on sait que les algorithmes « rêvent », au moins une partie du temps (cf. les phénomènes de « *spurious correlations* », de sur-apprentissage ou apophénie – c'est-à-dire de détection de « modèles » ou « patterns » qui correspondent à du bruit plutôt que des signaux – qui sont d'autant plus fréquents que les volumes de données sont importants). Dans une approche « *harm-based* » de la protection de la vie privée, sur un plan descriptif, ma vie privée est mise à mal lorsque quelqu'un d'autre acquiert à mon propos un « savoir » nouveau, ou, à tout le moins, une croyance à mon propos qui est justifiée. La question dès-lors se pose de savoir si de « simples » inférences produites par des algorithmes sur base de corrélations établies suivant des logiques opaques (*black box*) constituent un savoir ou une croyance suffisamment justifiée pour constituer une perte de vie privée. Jeffrey Skopek²⁶ répond par la négative, comparant les inférences algorithmiques aux rêves de quelqu'un d'autre, lesquels, quelles que soient les apparitions que l'on puisse y faire, ne peuvent s'interpréter comme des atteintes à notre vie privée. En quoi, effectivement, les « apparitions en rêve » sont-elles plus attentatoires à la vie privée lorsque ce sont des machines, plutôt que des humains, qui rêvent? On conçoit bien que dans une approche plus européenne de la protection de la vie privée comme attribut et instrument de protection de l'autonomie individuelle, l'argument puisse choquer. Il a pourtant le mérite de nous faire entrevoir que si les inférences (« vraies » ou « fausses ») algorithmiques posent aucun problème – et elles posent des problèmes cruciaux dès-lors que l'on se fonde dessus pour agir – ces problèmes ne sont peut-être pas idéalement appréhendables sous l'angle de la protection de la vie privée et des données personnelles. C'est dans le sillage de l'abandon de la causalité au profit de

²⁶ J. SKOPEK, « Privacy in Numbers? Biological Bodies of Data, Big Data's Epistemology, and the Legal and Ethical Status of Health Inferences », *2016 Annual Conference: Big Data, Health Law, and Bioethics*, Harvard Law School, 6 mai 2016.

la corrélation qu'il convient, me semble-t-il, de tenter de mieux identifier ce qui, spécifiquement, mérite protection. Il convient donc, pour bien saisir les enjeux, de nous attacher un instant à décrire les ambitions épistémiques sous-tendant le déploiement d'une rationalité algorithmique, à travers une brève description du phénomène des Big data, des algorithmes et des modes de « validation » de ceux-ci.

SECTION 1. – *Big Data*

Évoquer les *Big Data*, c'est évoquer d'emblée un changement d'approche dans la détection, la classification, l'évaluation anticipative des événements du monde et des comportements et propensions de ses habitants, c'est-à-dire, donc, une nouvelle manière de rendre le monde « prévisible » à défaut de le rendre « signifiant » (se passant des processus d'énonciation et de véridiction classiques) articulée à de nouveaux modes d'exercice du pouvoir : une nouvelle « gouvernementalité »²⁷. Dans la mesure où l'« intelligence des données », ravivant une sorte de comportementalisme numérique, supplanterait progressivement les formes – statistiques, politiques, juridiques, ... – à travers lesquelles nous nous représentons le réel, il convient de se demander comment le Droit serait encore en mesure de contenir, de borner, de limiter l'emprise d'une gouvernementalité algorithmique, y compris sur les processus législatifs et judiciaires. Le droit, court-circuité dans sa fonction anthropologique, n'est-il plus rien d'autre – dans le monde des robots, des algorithmes – qu'une contrainte parmi d'autres ?

Évoquer les *Big Data* c'est aussi évoquer d'emblée de nouvelles perspectives d'innovation technologique, de nouveaux services, de plus en plus personnalisés, capables d'anticiper plutôt que de seulement réagir aux stimuli du monde numérisé. Il ne s'agit plus tant de prévoir en vue de prévenir que de détecter dans l'actuel de pures potentialités et d'agir « par avance » comme si celles-ci étaient « réalisées » ou « actualisées ». Il ne s'agit donc plus tant de « réagir » à des « stimuli » du monde que d'anticiper les événements du monde en produisant les stimuli adéquats. La justification formelle des récoltes massives de données sur les plateformes, sont, bien souvent, l'« amélioration de l'expérience de l'utilisateur » : un certain type d'expérience de l'utilisateur qui rencontre les exigences de l'actionnariat des plateformes, expérience à travers laquelle l'utilisateur

²⁷ À propos de la notion de gouvernementalité, voy. M. FOUCAULT, « La gouvernementalité », in *Dits et écrits*, t. II, Quarto, Paris, Gallimard, 1994, 635-657 ; G. BURCHELL, C. GORDON et P. MILLER (eds), *The Foucault Effect. Studies in Governmentality*, U. Chicago Press, 1991.

devient de plus en plus « producteur » de revenus pour les plateformes, en « cliquant » davantage sur les liens publicitaires qui lui sont proposés, en laissant davantage de « signaux » permettant de mieux le « profiler », etc.²⁸ L'« expérience utilisateur » idéale est celle où il n'a plus à faire de choix : la machine et l'environnement s'adaptent automatiquement au profil de l'utilisateur, sans que celui-ci ait encore à former ni à formuler aucune préférence ou intention. Tout au plus réagit-il, sur le mode reflex, aux stimuli numérique qui lui sont adressés : il « aime », il n'« aime pas », il « retweet », il « poke », il « clique » – et « apprend » ainsi aux algorithmes à mieux le profiler.

Accueillies par les « techno-optimistes » comme la victoire tant attendue de l'analyse rationnelle sur les biais et préjugés humains et par les « techno-pessimistes » comme signifiant irrémédiablement la fin de la vie privée et l'avènement de la surveillance de masse, le phénomène des *Big Data* signifie surtout le franchissement d'un seuil de quantité (soft data émanant des réseaux sociaux, des moteurs de recherche, des blogs ; hard data émanant des administrations publiques ; méta-données produites automatiquement par les machines à propos des données ; données proliférant de l'internet des objets ; ...), de complexité (textes, images, vidéos, trajectoires, localisations, ...), de rapidité de prolifération des données à partir duquel nous, êtres humains, serions *contraints* d'automatiser et d'accélérer (pour tenir compte de l'accroissement continu, à grande vitesse, des masses de données) les processus de transformation des données numériques en informations opérationnelles²⁹.

L'expression *Big Data* renvoie donc aux masses de données numériques complexes à accumulation rapide, mais aussi à l'ensemble des nouvelles techniques logicielles (*Data Mining*, *Machine Learning*, *Neural networks*, *Social Network Analysis*, *Predictive Analytics*, « *Sensemaking* », *Natural Language Processing*, *Visualization*, ...) sans lesquelles les données resteraient « muettes ».

SECTION 2. – Des machines apprenantes

Pour devenir capables de détecter les corrélations utiles (prédictives) dans les masses de données, les algorithmes doivent apprendre. L'apprentissage est dit « supervisé » lorsque l'algorithme est entraîné sur des données

²⁸ FR. PASQUALE, « Privacy, autonomy, and internet platforms », in *Privacy in the Modern Age. The Search for Solutions* (M. ROTENBERG et J. HORWITZ), The New Press, 2015, pp. 166-167.

²⁹ D. WEINBERGER, *Too Big to Know : Rethinking Knowledge Now That the Facts Aren't the Facts, Experts Are Everywhere, and the Smartest Person in the Room Is the Room*, New York, Basis Books, 2012.

d'apprentissage fournies par le superviseur humain, qui contiennent à la fois les données et les résultats attendus (par exemple : des paramètres médicaux et des diagnostics) de manière à permettre son fonctionnement autonome sur des jeux de données pour lesquels les résultats sont inconnus, dans un processus de généralisation. La supervision sert à valider et à (re)calibrer le modèle retenu par l'algorithme (ce qu'il aura détecté comme « bonne » solution) de manière à aider le système à orienter ses modélisations dans la direction désirée.

L'apprentissage est dit « non supervisé », ou *bottom-up*, lorsqu'on ne fournit pas au système de modèle connu *a priori*. Aucun jeu de données d'entraînement ne lui est fourni, aucune « bonne solution » ne leur est présentée en modèle. On laisse alors l'algorithme analyser les données et identifier des corrélations entre celles-ci dans l'espoir de voir apparaître des modèles sous-jacents. Un exemple d'un tel algorithme est celui de groupement, qui permet de faire émerger au sein d'une population des individus « similaires ». L'algorithme auto-apprenant est capable de produire des solutions inattendues, des *patterns* ou modèles radicalement neufs, imperceptibles aux sens ordinaires et, en particulier, à l'œil humain³⁰.

SECTION 3. – Quelle validation des dispositifs apprenants ?

La mise à l'épreuve et la validation de ces dispositifs n'est pas chose aisée. La répliquabilité des opérations algorithmiques, par exemple, est quasiment impossible dans un contexte dans lequel les jeux de données impliqués sont en expansion continue. D'ailleurs ces modélisations dont on peut dire qu'elles sont produites à même le monde numérisé plutôt qu'à propos du monde physique ne visent aucunement à décrire la « vérité » mais seulement à être « opérationnelles ». La question de la validité ne se pose plus en termes de « vérité » mais de « fiabilité » – « *reliability without truth* », écrit Eric Winsberg³¹, une fiabilité réputée d'autant plus grande que les processus sont automatiques et évitent l'intervention humaine – et la dispense de la recherche de la vérité, de même que de l'historicité et de la causalité, est précisément l'un des moteurs de la nouvelle rationalité algorithmique. C'est l'idée de la « boîte noire » : on sait ce qui « entre » d'un côté, on constate ce qui sort de l'autre côté, mais

on ne sait pas ce qui se passe entre les deux. Que ce qui advient dans le monde ne se conforme pas au modèle produit algorithmiquement, au réel algorithmique – c'est-à-dire, lorsque ce qui arrive dans le monde fait « mentir » le profilage qui en avait été fait – n'est nullement un échec, ni un *raté* : ces notions d'échec, de *raté*, n'ont pas de sens dans une réalité numérique où tout écart par rapport à un modèle statistique, tout faux positif ou faux négatif, pour autant que ceux-ci soient détectables, est immédiatement assimilé dans la base statistique pour servir à affiner le modèle. C'est le principe même de l'apprentissage des machines (*machine learning*), supervisé ou non supervisé.

Indépendamment des questions relatives à la validation des mécanismes d'apprentissage, la fiabilité du « savoir » produit par l'analyse des *Big Data* est tout sauf assurée, pour d'autres raisons encore.

Premièrement, ce n'est pas parce qu'elles ont l'air de « se récolter toutes seules » (les données proliférantes du monde numérique – qu'elles émanent, comme des phéromones, des trajectoires et comportements, qu'elles soient produites par les administrations publiques ou qu'elles soient produites automatiquement par les machines (métadonnées)) que les données sont pour autant adéquates et exactes.

C'est que la qualité des données et leur adéquation dépend très fortement de la qualité et de la disposition des « capteurs ». On sait, par exemple, que les « gadgets » du *quantified self*, comme les « bracelets connectés » de différentes marques produisent des mesures différentes pour une même personne ou que le système *Street Bump* – un système de *crowdsourcing* permettant aux habitants d'une ville de renseigner les autorités locales à propos de l'état des routes en leur envoyant automatiquement les signaux GPS de leurs trajectoires (celles-ci traduisant l'état des routes, suggérant les obstacles, les trous dans la chaussée etc.) – est beaucoup plus fiable dans les zones fréquentées par une population nantie équipée de Smartphones performants que dans les zones moins affluantes, dont la population est moins bien équipée.

La qualité et l'adéquation des données dépend en outre de la « disposition » des informations pertinentes à se laisser numériser : si l'information cruciale pour la formulation d'une bonne politique publique, par exemple, n'était pas collectable par les capteurs mais se trouvait plutôt « entre » les qualités numérisables du monde et notre perception physique, corporelle (nos affects) de cette qualité du monde ? Alors, la « démondanisation » opérée par la numérisation empêcherait la transmission de cette information cruciale aux décideurs politiques. Par ailleurs, des logiciels de reconnaissance faciale pourraient très bien avoir plus de facilité à reconnaître les visages de caucasiens que les visages d'africains, en raison de la

³⁰ M. HILDEBRANDT, *Smart Technologies and the End(s) of Law*, Edward Elgar, 2015, p. 24.

³¹ E. WINSBERG, « Models of Success versus the Success of Models : Reliability without Truth », *Synthese*, septembre 2006, vol. 152, issue 1, pp. 1-19.

plus grande quantité d'images, sur Internet, représentant des personnes blanches que d'images représentant des personnes noires... De même il circule beaucoup moins d'images de femmes ayant plus de cinquante ans que de femmes plus jeunes et d'hommes de tous âges sur Internet, ce qui laisse préjuger d'une reconnaissance automatique des visages plus efficace pour les hommes blancs.

Enfin, que les modélisations algorithmiques produisent des résultats « vertueux » ou « diaboliques » est tributaire, aussi, du fait que les algorithmes apprennent *de nos propres comportements* (et nous ne sommes pas des anges !), ceux-là même dont transpirent les données numériques qui les « entraînent » et les « nourrissent ». Ainsi, par exemple, le « chatbot » de Microsoft prénommé Tay, une intelligence artificielle conçue pour converser en temps réel avec les utilisateurs de Twitter, apprenant par imitation des conversations ayant cours sur le réseau social, s'est-il, moins de 24 heures après son arrivée sur le réseau social le 23 mars 2016, mis à proférer des insultes racistes et sexistes. Les algorithmes, par leur seule automaticité, sont incapables de purger le monde social de ses travers, dans la mesure, précisément, où ils « métabolisent » automatiquement – c'est-à-dire en évitant de les évaluer préalablement – les données qui transpirent d'un état de fait, c'est-à-dire d'un monde façonné par les normes sociales et idéologies qui le traversent.

Les systèmes d'aide au recrutement dans le domaine de la gestion des ressources humaines proposent, sur base d'algorithmes entraînés sur les données relatives aux employés émanant des employeurs de tel ou tel secteur dans telle ou telle région, de « hiérarchiser » les candidats en fonction de « scores » censés attester de leur potentiel professionnel. Mais rien ne garantit que les données qui ont servi à l'apprentissage ne sont pas le reflet de pratiques discriminatoires défavorisant systématiquement les femmes, les personnes de telle ou telle origine ethnique, etc. L'algorithme entraîné avec des données reflétant certaines normativités sociales discriminatoires reproduira celles-ci aveuglément tout en les rendant moins perceptibles et contestables.

Autre exemple, encore, la « robotisation » de la justice (*smart administration of justice*) est parfois évoquée comme une solution aux problèmes de corruption de la magistrature dans certains pays africains, mais les algorithmes judiciaires ne feront jamais que reproduire ce qu'ils auront appris en métabolisant la jurisprudence passée, c'est-à-dire le résultat de l'action des juges corrompus qu'ils sont censés remplacer tout en rendant plus difficile la contestation des décisions machiniques.

De plus, indépendamment des méthodes d'apprentissage, les phénomènes d'*apophénie* ou de sur-apprentissage sont d'autant plus probables

que les quantités de données à analyser sont importantes³². C'est-à-dire, que les algorithmes auront une propension à identifier des modèles dans ce qui est en réalité du bruit statistique³³, un peu comme les étranges formes animalières que prennent les nuages dès lors qu'on les regarde un peu longuement dans un état de rêverie ou de somnolence. La puissance de calcul des machines dites intelligentes leur confère les capacités de l'entendement, – notion que Kant décrivait comme la faculté de synthétiser du divers – mais ne leur donne pas pour autant la raison, c'est-à-dire la faculté de donner un sens aux formes synthétisées. Le sens est produit par les concepteurs humains, en amont, notamment lorsqu'ils « nettoient » les données, ou lorsqu'ils décident des « métriques », des seuils d'intérêt, etc. qui, permettant de distinguer le « signal » du « bruit », bornant ainsi l'activité combinatoire des calculateurs.

Les « faux positifs » on le conçoit, ne sont pas tellement problématiques dans les applications de « *smarter marketing* » ... au pire, l'algorithme enverra des publicités pour des voitures à une personne qui n'a pas de permis de conduire. Les faux positifs ne sont pas tellement problématiques non plus dans le contexte de d'applications visant à écarter les courriels non sollicités (spams) de la masse proliférante de ce qui nous somme constamment de répondre. Ils sont d'avantage problématiques lorsque les processus algorithmiques aboutissent à priver quelqu'un de sa liberté ou de sa vie, l'empêcher d'obtenir un prêt ou d'accéder à un emploi.

En avril 2014, au cours d'un symposium à l'Université John Hopkins, le Général Michael Hayden, qui avait été directeur de la CIA et de la NSA le disait explicitement : « On tue des gens sur la base de métadonnées ». Pas des Américains, mais des étrangers, a-t-il ensuite expliqué, histoire de rassurer l'audience. De quoi s'agissait-il ? Le programme SKYNET de la NSA, un projet de *machine learning* relativement standard. Afin de lui apprendre à reconnaître de potentiels terroristes, l'algorithme est nourri de données d'apprentissages : les métadonnées relatives à un nombre (très restreint) de terroristes avérés. N'ayant « en magasin » que trop peu de terroristes avérés pour entraîner et tester l'algorithme, le programme se serait servi, à la fois pour l'entraînement et pour la validation, du même « jeu » de données relatives à sept terroristes avérés, ce qui n'est bien évidemment pas une méthode valide scientifiquement. Ensuite, l'algorithme

³² C. S. CALUDE et G. LONGO, « The Deluge of Spurious Correlations in Big Data », *Foundations of Science*, 2017, 22 (3):595-612.

³³ A. BAR-HEN, « Apophénie, ou l'art de prendre des vessies pour des lanternes », 16 mai 2016, <http://statistique.blogs.sciencesetavenir.fr/archive/2016/05/12/apophenie-ou-l-art-de-prendre-des-vessies-pour-des-lanternes-23485.html>

est « lâché » sur les métadonnées de tous les utilisateurs de téléphones mobiles du Pakistan à la recherche de « patterns » qui correspondraient à ceux qu'il a trouvés dans les données d'entraînement. L'algorithme s'est mis à produire une quantité importante de « faux positifs », désignant des personnes innocentes comme potentiels terroristes et cibles à abattre pour les drones, et qui ont été tuées par des frappes de missile Hellfire tirés par les drones Predator ou Reaper. Comble de l'ironie, sur une présentation power-point de la NSA révélée par Edward Snowden³⁴, la cible ayant obtenu le plus haut « score » de probabilité d'être terroriste n'est autre que le chef de bureau de la chaîne de télévision Al-Jazeera à Islamabad. « Faux positif » : le journaliste devait, pour les besoins de ses enquêtes et interviews, se rendre en des lieux – révélés par ses métadonnées téléphoniques – qui le rendaient suspect.

La croyance en l'objectivité des prédictions algorithmiques, en leur effectivité et en leur opérationnalité court-circuite bien souvent, chez ceux qui les adoptent à diverses fins (prévention de l'insécurité et du terrorisme, détection des propensions à la fraude, prédiction des comportements d'achat, optimisation des ressources humaines, ...), le processus d'évaluation critique de ce qui se présente, le plus souvent, comme une recommandation ou un système automatisé d'aide à la décision. Dans la mesure où ces dispositifs automatiques sont achetés et mis en service précisément pour accélérer et éviter la contestabilité des processus décisionnels, leurs « prédictions » se traduisent souvent quasi automatiquement en actions et interventions, lesquelles, à leur tour, modifient l'état des choses d'une manière qui ne permet plus d'identifier, contrefactuellement, ce qui se serait produit si la recommandation automatique n'avait pas été suivie.

Ainsi, l'anticipation ne fait-elle pas que décrire l'avenir, elle est « performative » : elle transforme l'avenir de manière telle qu'il devient extrêmement difficile – par manque de « ground truths » – de « mettre à l'épreuve » les algorithmes auto-apprenants pour évaluer effectivement leur validité. Cette perte de contestabilité³⁵ signifie, « tant pour les individus qui font l'objet de profilages que pour ceux qui se fondent sur ces profilages pour prendre des décisions à l'égard des individus », un déclin de la responsabilité – une raréfaction des occasions de « répondre » –, allant d'une dispense à une impossibilité de rendre compte des raisons de ses propres comportements et décisions.

³⁴ <https://theintercept.com/document/2015/05/08/skynet-courier/>

³⁵ A. ROUVROY, « The end(s) of critique : data-behaviourism vs. Due process », in *Privacy, Due Process and the Computational Turn* (M. HILDEBRANDT et E. DE VRIES eds), Routledge, 2012. Available at : http://works.bepress.com/antoinette_rouvroy/44.

Face à ce phénomène, deux voies de solution se profilent : la première – fondée sur l'assomption d'une superposition parfaite entre l'objectivité machinique/la vérité et la justice – consiste à s'assurer techniquement de l'objectivité, du caractère non biaisé, des modélisations algorithmiques (*auditing d'algorithmes*, ...).³⁶ La difficulté étant, notamment (outre l'opacité résultant des obstacles juridiques à l'exigence de dévoilement des secrets industriels ou des secrets défense), que la « logique » des traitements n'est pas nécessairement traduisible sous une forme linéaire communicable aux êtres humains. Ce n'est pas gratuitement d'ailleurs, que les ingénieurs impliqués dans la « fabrication » des algorithmes recourent à toute une série de métaphores biologiques et animalières pour désigner leurs algorithmes³⁷.

La seconde voie – fondée sur l'assomption d'une différence/différence fondamentale entre les idéaux d'objectivité et de justice – consiste à exiger la justiciabilité des décisions³⁸ affectant les individus, quelles soient – ou non – fondées sur des traitements automatisés de données. Il ne s'agit là plus tant de répondre de l'objectivité des processus algorithmiques que du caractère juste, équitable, légitime, des décisions prises, qu'elles soient ou non fondées sur ces processus. Autrement dit, il s'agit de faire ré-émerger de la « non-nécessité » – sans laquelle il n'est pas de décision (décider réellement présuppose qu'aucune solution ne s'impose par la nécessité), mais seulement de l'obéissance ou du conformisme –, de faire droit à l'incalculable,

³⁶ Voy. par exemple l'étude menée par A. DATTA, M. C. TSCHANTZ et A. DATTA, « Automated Experiments on Ad Privacy Settings A Tale of Opacity, Choice, and Discrimination », in *Proceedings on Privacy Enhancing Technologies* 2015 (1):92–112 <http://www.andrew.cmu.edu/user/danupam/dtd-pets15.pdf>. Dans cette étude, les auteurs mettent par exemple en évidence que le résultat de l'interaction entre les comportements des usagers, leur déclaration d'être de sexe masculin ou féminin dans préférences déclarées dans leurs paramètres d'« Add Settings » et les logiciels de personnalisation des messages publicitaires de Google est que les personnes s'étant déclarées de sexe masculin ont une beaucoup plus grande probabilité que celles qui se sont déclarées de sexe féminin de se voir envoyer des publicités relatives à des professions fortement rémunératrices (publicités pour un coaching de cadres supérieurs).

³⁷ Autant dire d'emblée, dès lors, que lorsque le RGPD (Règlement général sur la protection des données), dans les cas de décision fondée sur un traitement automatisé de données, prescrit la possibilité d'exiger la « communication des logiques de traitement » nous avons là une belle démonstration de la réticence, pour les législateurs, à envisager le mode d'existence spécifique des agents numériques évoluant dans le monde des données massives. Ces logiques de traitement sont pour nous, par définition pourrait-on dire, inintelligibles. Tout au plus peut-on « auditer » les « boîtes noires » que sont les algorithmes : c'est-à-dire les faire « tourner » sur différents jeux de données pour observer les différences dans les résultats et inférer sur cette base à certains aspects du fonctionnement de l'algorithme.

³⁸ Dans cette veine, lire not. D. KEATS CITRON et Fr. PASQUALE, « The Scored Society : Due Process for Automated Predictions », *Washington Law Review*, 2014, vol. 89, 2014, p. 1 ; *U. of Maryland Legal Studies Research Paper No.* 2014-8. Available at SSRN : <http://ssrn.com/abstract=2376209>.

à l'indécidable par le calcul. Il conviendrait d'identifier de quelle manière, par exemple, l'inversion de la charge de la preuve dans les cas de suspicion de discrimination indirecte induite par des recommandations automatisées, ou encore l'instauration d'un principe général de justiciabilité des décisions prises sur la base de traitements automatisés pourraient contribuer à restituer aux individus leurs « capacités à rendre compte, à énoncer par eux-mêmes » de ce qui les fait agir ou décider de telle ou telle manière, cette capacité d'énonciation – avec toutes les capacités de fabulation qu'elle suppose – étant au cœur de la notion de subjectivité juridique peut-être même d'avantage que les capacités d'entendement et de volonté classiquement tenues pour occuper le centre de gravité du sujet de droit.

Il est en tous cas indispensable, quel que soit le mode de « coopération » entre Humains et machines que puissions vouloir, dans quelque secteur d'activité ou de gouvernement que ce soit, de bien mesurer au préalable les spécificités des modes d'existence respectifs des Humains et des machines. L'exploration attentive de ces spécificités devrait notamment permettre de « dégonfler » radicalement les prétentions du « solutionisme technologique » à résoudre tous les problèmes, et d'identifier précisément, en évitant les pièges des perspectives paranoïaques et technophobes, ce qui nécessite notre vigilance.

Il convient donc à présent d'observer un peu plus précisément « comment » les algorithmes « regardent » le monde social et en quoi leur « regard » est différent du regard de la statistique « classique ».

CHAPITRE 3. Sous le regard des algorithmes : une « dissociété » ?

SECTION 1. – Une modélisation « immanente » au social numérisé ?

Une première caractéristique des traitements de données de type *Big Data*, consiste en ce que, plutôt que de subsumer les données dans des catégories préconstituées (comme les catégories statistiques, dont la construction suppose un processus conventionnel qui peut prendre du temps, ou les catégories sociales dans lesquelles les individus se reconnaissent), les « traitements de type *Big Data* » consistent à « faire surgir les "catégories" de la masse des données elles-mêmes », quasiment en temps réel. Ces « catégories » algorithmiques, que l'on appelle aussi *patterns*, ou modèles prédictifs,

ou encore profils (lorsqu'elles concernent les comportements humains), sont des « motifs » (on parle d'ailleurs de « visualisation » des données) évolutifs, formés par les corrélations observées non plus dans le monde physique mais entre des données numériques recueillies dans des contextes hétérogènes, indépendamment de toute explication causale. Pour le dire autrement, à la différence des traitements statistiques classiques, dans lesquelles les hypothèses ou catégories statistiques précèdent et président à la collecte des données, dans les traitements de type *Big Data*, c'est exactement l'inverse qui se produit : « la prolifération (la collecte) et le traitement des données précèdent et la détection des hypothèses ou catégories ».

Cependant, l'« objectivité » de la constitution algorithmique des profils, ne visant plus *a priori* personne en particulier, ne présupposant plus aucune catégorie perceptuelle, et en cela parfaitement « égalitaires », font aussi des modélisations et classifications fondées sur les *Big Data* un phénomène en apparence indépendant des systèmes de différenciations juridiques ou traditionnelles (en fonction du statut, de privilèges, d'avantages ou désavantages socio-économiques...) identifiés par Boltanski et Thévenot comme ce sur quoi s'appuie un modèle de cité qui en justifie ou en légitime les « états de grandeur » et dont l'existence est à la fois une condition et un effet des relations de pouvoir³⁹. De même, alors que le droit européen, notamment, reconnaît et protège une série de caractéristiques (le sexe, l'orientation sexuelle, le handicap, l'âge, l'origine ethnique, l'origine nationale, et la religion ou les convictions) particulièrement susceptibles d'exposer ceux qui en sont porteurs à des phénomènes de discrimination, les discriminations éventuellement susceptibles d'émerger dans le monde des *Big Data* seront probablement très difficilement « rattachables » (du moins directement) à ces différentes caractéristiques. En d'autres termes, alors que, par exemple, les directives européennes en la matière visent à prévenir les distinctions de traitement fondées sur la race ou l'origine ethnique, la religion ou les convictions, le handicap, l'âge, le genre ou encore l'orientation sexuelle, cette approche « catégorielle » de la discrimination semble *a priori* exclure les distinctions de traitement fondées sur l'« intelligence des données », et qui, comme les « machines a-signifiantes » décrites il y a près de trente ans par Félix Guattari, « ne connaissent ni les sujets, ni les personnes, ni les rôles, ni même les objets délimités. C'est précisément ce qui leur confère une sorte de toute-puissance, elles passent à travers les systèmes de signification au sein desquels se reconnaissent et s'aliènent les sujets individués »⁴⁰.

³⁹ L. BOLTANSKI et L. THÉVENOT, *De la justification. Les économies de la grandeur*, Paris, Gallimard, 1991, p. 162.

⁴⁰ F. GUATTARI, *Révolution moléculaire*, coll. Encre, Recherches, 1977, p. 264.

Autrement dit, aux systèmes de perception et d'interprétation du monde précédemment fondés sur des phénomènes de représentation (représentation statistique, témoignage langagier, symbolisation, institutionnalisation,...) et de reconnaissance de structures, formes, catégories (politiquement, juridiquement, culturellement) préconstituées, l'hyperfragmentation et la croissance exponentielle de l'univers numérique substitue la possibilité de modélisations simultanées plutôt qu'antécédentes aux traitements de données. Pour cette raison, la reconfiguration radicale de ce que Michel Foucault appelait les « pratiques divisantes »⁴¹ (présidant aux distinctions de traitement opérées entre les individus) par les traitements de type *Big Data* défie – plus fondamentalement encore que les régimes de protection de la vie privée et des données personnelles – le droit de la non-discrimination, lequel a toujours été pensé en fonction de la préexistence de catégories, de groupes humains préconstitués, clairement reconnaissables et à ce titre particulièrement vulnérables aux pratiques discriminatoires⁴².

SECTION 2. – Une modélisation des cas singuliers ou une « personnalisation » à l'échelle industrielle

Il en va de l'aura d'impartialité des algorithmes : il s'agit de ne plus inscrire les personnes dans aucune « catégorie *a priori* ». Comme l'explique Olivier Rey dans *Quand le monde s'est fait nombre*, dans la conceptualisation des rapports entre l'individu et le collectif, il fut une époque – avant le XVII^e siècle, où l'individu n'existait que comme émanation de sa communauté (communauté > individu). Cette représentation-là fut remplacée ensuite par une autre, dans laquelle les individus, s'émancipant de toute communauté originaire, sont premiers et acceptent volontairement de « faire société », sur un modèle très contractuel (individu > société). « L'"individu", de terme d'une division, se mit à être envisagé comme

⁴¹ « (...) j'ai étudié l'objectivation du sujet dans ce que j'appellerai des "pratiques divisantes". Le sujet est soit divisé à l'intérieur de lui-même, soit divisé des autres. Ce processus fait de lui un objet. Le partage entre le fou et l'homme sain d'esprit, le malade et l'individu en bonne santé, le criminel et le "gentil garçon" illustre cette tendance » (M. FOUCAULT, « Le sujet et le pouvoir », *Dits et Ecrits II*, Paris, Gallimard, 2001, p. 1042).

⁴² Nous verrons que les approches *Big Data* questionnent l'efficacité de l'approche actuelle consistant à certain improvements interdire et à empêcher les distinctions de traitement fondées sur des caractéristiques protégées ou catégories vulnérables d'une part, et en empêchent toute action collective des victimes potentielles de discriminations fondées sur le profilage algorithmique d'autre part.

point de départ, extérieur et antérieur au monde partagé, à partir duquel les réalités collectives se constituent. Pour le résumer en une formule : "Je" était le singulier de "nous", "nous" devient le pluriel de "je" »⁴³.

Il semble que le tournant computationnel marque une nouvelle étape dans l'histoire des rapports entre l'individu et le collectif, dans lequel aucune société, aucun contrat social, aucune règle de justice fût-elle collectivement consentie ne doive plus prévaloir sur l'adaptation, en temps réel, des « normes » aux « profils » des (in)dividus. Cette ultime hyper-individualisation passe par une modélisation des cas singuliers, ou une « personnalisation » à l'échelle industrielle.

Jamais les interactions entre la personne et les bureaucraties privées et publiques n'ont été à ce point individualisées grâce au profilage algorithmique. Ce phénomène d'individualisation n'a rien à voir avec la prise en compte de la singularité des personnes. Il s'agit seulement de substituer, aux catégorisations et qualifications *a priori* toujours trop générales et abstraites, politiquement débattues, idéologiquement contestables, culturellement biaisées, des profils impersonnels, opaques, implicites, indiscutables, c'est-à-dire des corrélations statistiquement prédictives entre variables comportementales impersonnelles détectées automatiquement dans la masse des données disponibles. L'indifférence à l'égard à la fois de ce qui fait non pas l'individualité mais la singularité d'une vie et de ce qui l'inscrit dans des contextes collectifs au profit de ce qui la relie statistiquement à des profils impersonnels prédictifs, est ce qui confère aux processus algorithmiques leur aura d'objectivité tout en les rendant, littéralement, injustifiables.

Cette étrange « sollicitude » impersonnelle pour l'individu présuppose une capture numérique qui passe à travers les masques sociaux, rôles ou positions assumés, au profit d'un principe de perméabilité de l'espace psychique individuel exposé à une « vision » algorithmique qui n'est plus limitée par aucun principe de séparation entre l'espace psychique et l'espace public (« vernis social », « pudeur », « politesse »,...), ni déformée ou informée par la *persona*, c'est-à-dire par le masque social – ce qui, de la personne, s'édite et s'adresse consciemment à autrui. Dispensant de toute comparution, le régime d'optimisation algorithmique n'est « pas un régime de vérité » qui obligerait les sujets à parler sur le mode de la confession, du témoignage, de l'aveu, mais « un régime d'indistinction entre les signaux numériques, assignifiants mais calculables, et le monde ».

Bref, c'en serait fini de nos « fictions » tout autant que de nos régimes d'administration des preuves et des mises à l'épreuve, puisque « nous

⁴³ O. REY, *Quand le monde s'est fait nombre*, coll. Les essais, Stock, 2016, p. 62.

n'aurions plus à rejoindre un monde dont nous serions séparés » : nous faisons, désormais, numériquement corps avec le monde numérique, immergés dans et traversés par les flux. Ce régime d'indistinction entre les signaux et les choses disqualifie la seule chose peut-être qui soit réellement imprévisible et incalculable à propos d'une personne : la manière toujours singulière suivant laquelle elle (se) rend compte d'elle-même et à elle-même, elle se rejoint dans ses propres gestes, dans « ce qu'il lui est arrivé (de faire) ». Si le sujet, comme le peuple, manque toujours à soi-même, si la seule liberté que nous ayons est de faire volontairement ce que nous voulons involontairement, l'énonciation « contrefactuelle », la rationalisation dans l'après-coup est le mode suivant lequel nous pouvons « être dignes de ce qui nous arrive »⁴⁴, suivant lequel la personne peut « donner une forme à son destin »⁴⁵.

§ 1. Le fantasme d'exhaustivité : une coextension de la « base statistique » au « réel numérisé » ?

Alors que les pratiques statistiques « classiques » éliminent des jeux de données tous les « points » de données qui seraient trop éloignés de la moyenne ou du « plus probable » comme générateurs d'erreurs et de perturbations, les traitements de type *Big Data* supposent au contraire de prendre « tout » en compte, y compris ce qu'il y a de plus ténu, de plus éloigné des grands nombres, ces points éloignés n'étant plus même rapportés à aucune moyenne (la notion même de « moyenne » perdant toute pertinence lorsque la « base statistique » est, comme dans le contexte des *Big Data*, hyper-plastique en extension). C'est que ce qui « compte » n'est plus la distance qui sépare chaque point de donnée d'une « moyenne », mais la force des corrélations qui relie entre elles des données disparates.

§ 2. À chacun son « coût » réel

C'est ce qui permet des processus de « personnalisation », c'est-à-dire la différenciation en fonction de « profils » (de criminels ou fraudeurs potentiels, de consommateurs, d'utilisateurs, ...) de plus en plus nombreux et précis, des interactions sécuritaires, commerciales, administratives, éducatives, médicales, ... Cela signifie notamment, en pratique qu'alors que dans le contexte des traitements statistiques traditionnels, on pouvait

dire que les statistiques valaient pour les grands nombres, pas pour les cas singuliers, l'approche *Big Data*, au contraire, vise une pertinence des « modèles » pour les cas les plus singuliers, ou, pour le dire plus simplement, vise à remplacer la catégorisation par la personnalisation.

Cela dit, le terme de « personnalisation » est trompeur dans la mesure où il pourrait faire croire que les profils reflètent la singularité des vies qui y sont reliées. Il faut ici bien comprendre que ce qui confère aux processus algorithmiques leur aura d'impartialité tout en les rendant, littéralement, « injusticiables », c'est précisément, d'une part, leur indifférence aux catégories et contextes collectifs dans lesquels vivent les individus (les processus algorithmiques y substituent des « profilages » qui ne se superposent explicitement à aucune catégorisation socialement éprouvée) et, d'autre part, le fait qu'en optimisant par avance les interactions entre l'individu et son milieu, ils court-circuitent ce qui fait la singularité non numérisable d'une vie, qui est, comme l'explique Judith Butler, « le fait même de pouvoir s'imaginer une vie, un temps au-delà du présent, une "forme" de vie »²¹.

Dans le domaine du crédit hypothécaire ou du crédit à la consommation par exemple, les applications des analyses de type *Big Data* permettent d'évaluer les risques de non remboursement propres aux individus en fonction, non plus de leur situation personnelle (situation professionnelle, financière, familiale, etc.), mais de la proportion statistique de mauvais payeurs résidant dans le même type de quartier, faisant leurs courses dans le même type de magasins, fréquentant tel ou tel type de forum de discussion sur Internet... Autant d'éléments *a priori* sans lien avec l'état de solvabilité actuel ou futur des individus mais ceux-ci se trouvent, statistiquement corrélés (sans que l'on sache pourquoi) à un taux de défaut de paiement supérieur à celui de la population générale. Au fond, de cette logique algorithmique purement inductive, émerge une image quasiment épidémiologique du crime, du défaut de paiement, comme quelque chose que les individus « attraperaient » comme par « contagion ». Il s'agit alors d'un déterminisme articulé, non aux caractères propres à l'individu, mais à ses trajectoires dans l'espace, aux relations qu'il entretient avec ses semblables, aux supermarchés qu'il fréquente, aux films qu'il regarde, aux musiques qu'il écoute, etc. La capacité subjective de « réticence », c'est-à-dire cette capacité que les individus ont de ne pas faire tout ce dont ils sont capables, de ne pas actualiser toutes leurs potentialités, d'où ils tiennent leur fondamentale imprévisibilité, se trouve alors circonvenue, contournée, par ce modèle épidémique, purement relationnel plutôt qu'intentionnel ou existentiel, du comportement humain saisi exclusivement à travers ses comportements et en fonction des opportunités et des coûts qu'il est susceptible d'actualiser.

⁴⁴ G. DELEUZE – dernier cours de Vincennes – Anti-oedipe et autres réflexions cours du 3 juin 1980.

⁴⁵ A. CAMUS, *Le mythe de Sisyphe. Essai sur l'absurde*, Paris, Gallimard, 1942.

Dans la société actuarielle, ou dans la société de l'assurance, la charge correspondant à la part incompressible d'incertitude radicale, qui découle de ce qu'on n'est jamais certains que tout ce qui est possible ou probable se réalise effectivement, est pris en charge par des formes diverses de mutualisation, c'est-à-dire de prise en charge collective, du risque⁴⁶.

Avec l'arrivée des *Big Data* (mais aussi des phénomènes de *quantified self*), cette mutualisation des risques tend à faire place à une approche beaucoup plus individualisante, tentant de déterminer, grâce notamment à la granularité des analyses de type *Big Data*, pour chacun, individuellement, ses « risques individuels » et ses « coûts réels » – une manière d'individualiser le risque et, du même coup, de détricoter les mécanismes de solidarité face à ce que l'on appelait autrefois « la providence ». Plutôt que de répartir la charge des risques, les politiques « préemptives » consistent à faire, anticipativement *comme si* l'événement redouté avait eu lieu et à prendre immédiatement, par avance, les mesures s'imposant en conséquence (refus d'assurance à un fraudeur potentiel, élimination préventive d'un terroriste potentiel, orientation professionnalisant des enfants sur base d'un profilage précoce, ...). Bien entendu, suivant les domaines d'application, la préemption peut être plus ou moins justifiée, pour autant que son acceptabilité fasse, *a minima*, l'objet d'un débat public.

Toujours est-il que la fascination pour les *Big Data* et les pratiques de datamining comme moyen d'imputer anticipativement à chacun – d'une manière à la fois très individualisante, singularisante, qui ne passe plus par aucun rapport à aucune « moyenne » calculée sur une population – les coûts et opportunités « réels » dont il est porteur (dans le domaine de la sécurité, de la santé, du marketing, des ressources humaines, de l'assurance, ...), n'est peut-être pas tant la conséquence d'un besoin accru de sécurité que d'une transformation de la réponse faite à une demande de sécurité qui, elle, n'est absolument pas neuve.

⁴⁶ Ainsi François Ewald expliquait-il, à propos du calcul des probabilités au cœur des pratiques des assurances, qu'il fonctionne comme « une ruse de la raison ». Le calcul des probabilités est un instrument d'investigation voué à pallier à l'impossibilité d'expliquer physiquement les phénomènes. C'est un instrument d'expérimentation par la raison pure. Ce n'est pas seulement que nous ignorions les lois régissant les phénomènes que nous percevons dans leur infinie variété et dans leur infinie dispersion, nous ignorons aussi leurs causes. Notre ignorance est telle que, si même nous inférons quelques régularités, nous serions incapables de déterminer si ces régularités seraient constitutives de lois. Le paradoxe du calcul des probabilités résulte du fait que cette ignorance fondamentale n'est pas vouée à être vaincue par aucun savoir qui proviendrait d'une découverte, que nous ne quitterions jamais le domaine de l'observation. Tout l'art du calcul consiste dès lors à faire jouer cette ignorance contre elle-même, de la neutraliser, pourrait-on dire, en l'utilisant contre elle-même (Fr. EWALD, *Histoire de l'État providence*, LGF – Livre de Poche, nouv. éd. ent. ref (1 janvier 1996), p. 114).

Dans *Le monde d'hier*, Stefan Zweig, situait dans les années 1900 à Vienne l'âge de la sécurité, l'âge d'or des assurances. À l'âge de la sécurité, on n'avait pas vu disparaître, comme par magie, toute espèce de danger, mais on avait appris, grâce à la statistique, à domestiquer, par le calcul des probabilités et la mutualisation des risques, ce qu'autrefois on eût appelé la providence et qu'on appellerait aujourd'hui l'incertitude.

« Si je cherche une formule commode qui résume l'époque antérieure à la première guerre mondiale, dans laquelle j'ai été élevé, j'espère avoir trouvé la plus expressive en disant : "C'était l'âge de la sécurité" ... Ce sentiment de sécurité était le trésor de millions d'êtres, leur idéal de vie commun, le plus digne d'efforts... peu à peu les grandes masses parvinrent à y accéder. Le siècle de la sécurité devint l'âge d'or des assurances... les domestiques prirent sur leurs économies une assurance vieillesse et payèrent à l'avance à la caisse mortuaire leur propre enterrement ».

L'individualisation « parfaite » des risques et opportunités signifierait par exemple la fin de la raison d'être des assurances, dont le rôle premier n'est certainement pas d'individualiser la charge des risques mais au contraire, de constituer des « contrats sociaux » restreints entre des personnes, les assurés, qui, soumis à des risques comparables, s'engagent à prendre en charge collectivement les coups du sort qui s'abattraient sur certains d'entre eux. Comme l'explique François Ewald,

« Il n'y a pas à proprement parler de risque individuel, sans quoi l'assurance se transformerait en gageure ou en pari. Ce n'est en effet que sur l'étendue d'une population que le risque devient calculable. Le travail de l'assureur est précisément de constituer cette population par sélection et division des risques »⁴⁷.

Par ailleurs, explique-t-il encore, l'assurance individualise, elle définit chacun comme risque, mais d'une individualité qui ne se réfère plus à une norme abstraite, d'une individualité relative aux autres membres de la population assurée, d'une individualité moyenne ou sociologique. Alors que l'accident, le dommage, la malchance et la souffrance sont toujours individuels, s'abattant sur un individu et épargnant un autre, le *risque* d'un accident, d'un dommage, d'une perte, ... affecte toujours – en tant que risque – une population. À proprement parler, il n'existe pas de risque individuel ; sans quoi l'assurance ne serait rien d'autre qu'un pari. Le risque ne devient quelque chose de calculable que lorsqu'il est réparti sur une population. Le travail de l'assureur consiste, précisément, à constituer cette population en sélectionnant et en divisant les risques. L'assurance

⁴⁷ *Ibid*, pp. 138-139.

ne peut jamais couvrir que des groupes ; elle fonctionne en socialisant les risques. Elle fait de chaque individu une portion du tout⁴⁸.

Au regard de ces considérations relatives à l'assurance, les *Big data* permettraient, semble-t-il, le passage d'une société actuarielle à une société post-actuarielle, dans laquelle la solidarité entre personnes relevant de la même « population » d'assurés, est remplacée par la possibilité d'une individualisation et de la fluctuation des primes en temps réel. L'adaptation, en temps réel et en continu, des primes dues par chaque assuré en fonction de son comportement journalier (qualité de la conduite automobile, pratique sportive, habitudes alimentaires, ...) pourrait avoir des effets d'incitation (ou de dissuasion) socialement souhaitables, mais pourrait aussi – en permettant par exemple d'évaluer l'état de santé futur des personnes, sur base, par exemple de l'historique de leurs achats de consommation courante – produire des conséquences néfastes sur le plan des principes d'égalité d'opportunités et de solidarité. À cet égard les enjeux de l'individualisation des primes (ou de l'hyper-segmentation du marché des assurances en fonction d'éléments de plus en plus individuels et singuliers) soulevés par le phénomène des *Big Data* ne sont pas foncièrement différents de ceux que soulèvent la mise à disposition des assureurs, employeurs et autres acteurs intéressés des données génétiques individuelles indicatrices de prédispositions à certaines maladies chez des individus actuellement en bonne santé⁴⁹.

SECTION 3. – Une civilisation algorithmique sans sujets et sans monde ?

§ 1. L'individu sans intérieur, somme des relations qui le constituent

Les profils, scores, rankings, qui adhèrent aux personnes plus sûrement qu'une seconde peau, au point de leur être coextensifs, nient tout écart, toute distinction entre la personne et ses comportements. L'individu hyper-indexé n'est plus « défini » par aucun état civil, par aucun « statut » social, par aucune « appartenance » (à aucun groupe social reconnaissable), ni « perçu » en fonction de sa sensibilité, son plaisir, sa peine

⁴⁸ FR. EWALD, « Insurance and Risk », in *The Foucault Effect : Studies in Governmentality* (G. BURCHELL, C. GORDON et P. MILLER eds), Chicago University Press, 1991, pp. 197-210.

⁴⁹ À cet égard, nous nous permettons de renvoyer le lecteur à notre étude sur le sujet. A. ROUVROY, *Human Genes and Neoliberal Governance: A Foucauldian Critique*, Routledge-Cavendish, 2007.

(en cela, le « régime » algorithmique se distingue fondamentalement des « régimes » utilitaristes), mais par la somme de ses « relations » (les données numériques reflètent toujours des relations, entre une personne et une autre personne, un objet connecté, un lieu, ...), relations ne relatant plus rien d'antécédent à elles-mêmes.

Obtenue dans un espace purement métrique, un espace neutralisé (par « neutralisation » il faut ici entendre le processus par lequel, « des régions alentours d'un existant deviennent de simples dimensions de l'espace »⁵⁰), la « décision » algorithmique en régime de comportementalisme numérique diffère radicalement de toute décision de justice. Elle nie la distinction, fondamentale pour le juge et pour la justice, non seulement entre la personne et ses actes⁵¹ mais aussi entre la personne est ses « propensions » ou puissances (les algorithmes identifient ce que les personnes sont « en puissance » : terroristes « en puissance », fraudeurs « en puissance », récidiviste « en puissance », ...). Au sujet de droit se substituent les profils algorithmiques, composant un moulage extrêmement plastique précédant chacun dans ses comportements, trajectoires, choix : un moule qui n'aurait plus rien de la fixité des traces mais se comporterait plutôt comme une empreinte qui précéderait les pas, une ombre jamais tout à fait adéquate qui aurait, sur les mouvements de la proie, toujours une longueur d'avance⁵².

§ 2. Il n'est pas de plus sûr moyen de se prémunir contre l'avenir que de le produire

Cette longueur d'avance, c'est l'espace spéculatif qu'ouvre la vision algorithmique, dans lequel elle agit par avance, sur ce que peuvent les corps. Ces visions oraculaires non figuratives et en ce sens dé-figurantes⁵³, visions d'opportunités et de dangers n'ayant vocation à être ni vraies ni fausses, n'existent qu'en vue d'immuniser l'actuel contre le virtuel, le probable contre l'excès du possible, l'impassibilité inorganique contre les

⁵⁰ FR. FISCHBACH, *La privation de monde. Temps, espace et capital*, Vrin, 2011, p. 58.

⁵¹ Selon P. Legendre, « [i]l n'y a pas d'autre justification aux procès intentés contre les assassins que celle-là : séparer de son crime celui qui tue, faire que sa part maudite devienne sa part de sacrifice. Cela s'appelle juger » (*La Fabrique de l'homme occidental*, G. CAILLAT, P. LEGENDRE et P.-O. BARDET, Arte, 1996, 80 min.).

⁵² Nous nous permettons de renvoyer ici à A. ROUVROY, « La vie n'est pas donnée », *Études digitales. Le gouvernement des données*, Classiques Garnier, 2016, n° 2, pp. 195-217.

⁵³ « Déchirer la représentation, le tissu de la transcendance, (...) est la condition nécessaire pour faire face (sans pourtant jamais faire face) à l'inimaginable » (B. MANCHEV, « Persistance de l'image et devenir-sensible du sensible. Georges Bataille et la surcritique de la représentation », *Le Portique*, n° 29, 2012, document 9).

événements organiques. La gouvernementalité algorithmique se présente donc comme un système immunitaire de la réalité numérique contre toute hétérogénéité incalculable, contre toute pensée du dehors⁵⁴ inassimilable, irréductible, non monnayable, non finalisée, inorganisable, souveraine⁵⁵, c'est-à-dire, aussi, *contre le monde* (ou contre l'univers qu'évoquait Asimov).

La gouvernementalité algorithmique n'est donc pas un régime de « normalisation », c'est un régime de « neutralisation », qui opère hors les normes, une régulation immanente, une radicalisation de l'ambition manifestée par la « gouvernance par les nombres », de liquider toute espèce d'hétéronomie⁵⁶. Congédiant le monde et ses représentations au profit d'une réalité faite de flux numériques⁵⁷ aux joints étanches indemnes d'événements elle neutralise la puissance des sujets : leur capacité d'énonciation contrefactuelle (qui est aussi capacité d'effacer ses traces, de mentir, de prétendre), leur capacité de réticence (de ne pas faire tout ce dont ils sont capables), leur capacité de décider (de trancher sur fond d'indécidabilité plutôt que d'obéir aux résultats d'un calcul), et leur capacité de projeter collectivement des possibles (dont dispense l'hyper-individualisation déjà évoquée).

Conclusion : Peut-on (se) gouverner hors les normes sans se dissoudre dans les flux ?

L'oubli des sujets derrière les profils est aussi une abstraction qui court-circuite la fonction anthropologique du droit. « Faire de chacun de nous un "*homo juridicus*" est la manière occidentale de lier les dimensions biologique et symbolique constitutives de l'être humain », écrit Alain Supiot⁵⁸. Si la « gouvernementalité algorithmique » disqualifie le sujet de droit, éclipçant ses dimensions biologique et symbolique au profit d'une

⁵⁴ M. FOUCAULT, « La pensée du dehors », *Dits et écrits*, t. I, 1954-1975, 1966, texte n° 38.

⁵⁵ J. FEYEL, « Le corps hétérogène de Georges Bataille », in *Actes du colloque international Projections : des organes hors du corps* (13-14 octobre 2006), <http://www.epistemocritique.org/IMG/pdf/ProjectionsFeyel.pdf>

⁵⁶ A. SUPIOT, *La gouvernance par les nombres. Cours au Collège de France (2012-2014)*, coll. Poids et mesures du monde, Paris, Fayard, 2015, p. 243.

⁵⁷ « (...) soit l'abolition de toute substance-sujet et de toute substance-objet : il n'y a que des ponctualités subjectives et objectives, des pauses momentanées dans la production indéfinie des flux » (Fr. NEYRAT, *Clinamen. Flux, absolu et loi spirale*, ère, 2011, p. 25).

⁵⁸ A. SUPIOT, *Homo juridicus. Essai sur la fonction anthropologique du droit*, coll. Points, Paris, Éditions du Seuil, 2005, p. 10.

appréhension numérique des (in)dividus, c'est au nom d'un « régime de vérité »⁵⁹ numérique que l'on voudrait émanciper du joug de la représentation, un régime d'immanence parfaite dans lequel les données parleraient d'elles-mêmes, dans une langue émancipée du langage lui-même, sans qu'il soit encore nécessaire de rien transcrire ni interpréter. C'est une certaine « passion du réel » qui réclame, en lieu en place du gouvernement politique et juridique, une forme de gouvernement cybernétique.

Les modélisations/optimisations anticipatives promettent de congédier le droit lui-même, quand bien même celui-ci serait à la source du lien social comme « forme excédant celle de la multiplicité des petits liens par lesquels les individus s'obligent de manière mutuelle ou se lient à l'univers des choses qui les entourent »²³. Enfin inséparés de l'univers dans un monde numérique qui dispenserait du monde, nous n'aurions plus besoin d'aucune forme excédant cette « multiplicité des petits liens » puisque ceux-ci s'agenceraient automatiquement, par anticipation sur toute effectuation consciente, dans les profils algorithmiques à travers lesquels les conduites pourraient être gouvernées par avance. La préemption algorithmique – organisant un espace parfaitement lisse et continu, exempt d'aspérités humaines, une temporalité unidimensionnelle – dispenserait à la fois de la désobéissance, du jugement et de la sanction, tout en rendant « les normes » à la fois indétectables en tant que synthèses algorithmiques d'états de faits, et incontestables.

Comment, face à cette liquidation algorithmique des formes à travers lesquelles nous (nous) gouvernons, retenir ouverte – comme un accroc dans la trame fluide du réel numérique – la possibilité du « projet » politique et la survivance des sujets ? Comment faire en sorte que les animaux hétérochroniques (hétérochroniques en ce sens que nous avons la capacité de n'être pas totalement « contenus » dans aucun présent, que nous sommes hantés par le passé et nous projetons dans le futur, et que c'est souvent dans l'après-coup que nous forgeons les motivations de nos actions) que nous sommes ne soient pas pris en compte seulement en tant qu'agréats temporaires de données exploitables en masse ? Comment, face à l'hyper-individualisme optimisateur, revitaliser l'espace public comme espace de délibération à propos de la chose commune irréductible à la seule concurrence des intérêts individuels et comme lieu d'exposition des limites de la représentabilité, c'est à dire en tant que lieu d'ouverture de nouvelles possibilités politiques ? Dans ce « réel inséparé » vers lequel

⁵⁹ Nous nous permettons de renvoyer à A. ROUVROY et B. STIEGLER, « Le régime de vérité numérique. La gouvernementalité algorithmique à un nouvel état de droit », *Socio. Le tournant numérique...et après ?*, 2015, n° 4, pp. 105-132.

nous ferions mouvement, le droit n'aura-t-il pas perdu définitivement sa fonction anthropologique ? Ne sera-t-il plus qu'une contrainte parmi d'autres ? Peut-on se gouverner hors les normes sans se dissoudre dans les flux ? Voilà quelques-unes des questions fondamentales auxquelles nous sommes à présent confrontés.